DOCUMENT RESUME

ED 311 028 SP 031 488

AUTHOR Salzman, Stephanie A.; Whitfield, Patricia T.
TITLE Teacher Competency Testing: Panacea or Pandora's Box?
PUB DATE Aug 89
NOTE 27p.; Paper presented at the Meeting of the Northwest Association of Teacher Educators (Tacoma, WA, July 30-August 2, 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Minimum Competency Testing; *Standardized Tests; State Standards; *Teacher Certification; Teacher Evaluation; *Test Bias; *Test Validity

ABSTRACT

This paper examines the major issues to be faced and the decisions to be made in the development and administration of basic competency tests used for the certification of teachers. The focus is on the use of competency tests as criteria in determining the initial certification of teachers, as differentiated from the ongoing observation and evaluation of practicing teachers. The first section of the paper addresses issues surrounding the definition of teaching competence and the resulting constructs to be assessed. The second section, on validation strategies, examines technical issues involved in teacher competency test validation and the relevance of different types of validity evidence. The final section provides an overview of issues related to equity and fairness of the testing process and the accompanying legal and ethical considerations. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED311028

Teacher Competency Testing:

Panacea or Pandora's Box?

Stephanie A. Salzman

Idaho State University

and

Patricia T. Whitfield

Idaho State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Salzman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented to the Northwest Association of Teacher Educators
July 30-August 2, 1989, University of Puget Sound, Tacoma, WA

BEST COPY AVAILABLE

SP031488

Teacher Competency Testing:

Panacea or Pandora's Box?

In response to the demand for improved teacher quality and increased accountability, many states and school districts have mandated policies requiring successful performance on competency examinations prior to teacher certification. Every state except Alaska and Iowa has adopted or is in the process of adopting some form of teacher testing program (Rudner, 1987). As of April 1987, 44 states had implemented or decided to initiate a program requiring prospective teachers to pass a written and/or performance test before being awarded full certification. Furthermore, 27 states were testing or planning to test applicants for admission to college or university teacher education programs.

The stated goal of evaluation of teachers at the preservice level is to predict the effectiveness of future teachers in the classroom in order to control the quality of individuals entering the profession (Murray, 1986). Test-enforced standards for admission to teacher preparation programs and for certification are seen by proponents as means to screen out unqualified individuals, to strengthen the profession, and to attract better qualified candidates. As a result of these competency testing programs, the public's confidence in teachers, teaching, and the schools is expected to improve (Sandefur, 1984).

While virtually all the states have made some form of commitment to teacher testing, the diversity of philosophies and attitudes toward the issue is substantial. The states differ in terms of when they test prospective teachers, what their tests cover, the difficulty of their

instruments, and which tests are used. The issue of teacher testing covers not only whether teachers should be tested, but also how they should be tested and when. Virtually every aspect of teacher testing has been subject to considerable debate.

Thus, while the number of testing programs is increasing, the attitudes and rationales behind them are evolving and their particulars are changing. Passing scores are being adjusted, plans are being altered, some programs are being abolished, and new programs are being established. In weighing their options, policymakers are confronted with the responsibility of trading emotional and financial costs against the benefits the testing programs may provide.

Meanwhile, colleges, universities, and departments of education using competency examinations as criteria for initial teacher certification are exerting much effort and expense to validate the tests, establish cut-off scores, arrange test-taking procedures, and inform and monitor examinees. In summarizing teacher evaluation, Sweeney and Maratt (1986) concluded, "We have an important activity beset by conflicting purposes and political infighting. We have uncertainty as to who should gather evidence on elusive criteria on which there has been more disagreement than agreement, plus a tangled web of measurement devices" (p. 453). The problems inherent in using competency examinations for teacher certification may be adequately addressed given enough time, hard work, and dialogue. Toward that end, empirical evidence must be considered in conjunction with practical judgments and the issues and requirements that characterize individual teacher education programs and certification agencies.

This paper examines the major issues to be faced and the decisions to be made in the development and administration of basic competency tests used for the certification of teachers. The focus of the paper is on the use of competency tests as criteria in determining the initial certification of teachers, as differentiated from the ongoing observation and evaluation of practicing teachers. The paper contains three major sections. The section titled Selection of Test Content addresses issues surrounding the definition of teaching competence and the resulting constructs to be assessed. The Validation Strategies section examines technical issues involved in teacher competency test validation and the relative relevance of different types of validity evidence. The final section, Standardization, provides an overview of issues related to equity and fairness of the testing process and the accompanying legal and ethical considerations.

Selection of Test Content

Any test of knowledge in professional and general education is bound to measure only a sample of the important qualities necessary to be a teacher. The selection of the particular content to be tested, therefore, becomes as issue in the construction and use of teacher competency examinations. A related issue is the appropriateness of performance assessments versus multiple-choice subject matter examinations as measures of teaching competence.

Since the early 1980's more than 275 education task forces have been organized in the United States, at least 18 books or book-length national reports have been generated, all intended to "fix" the schools (Orlich, 1985). Among the outcomes of these concerted efforts have been competency

tests in more than 40 states and changes in teacher certification requirements in nearly as many. Applicants for teacher certification are tested in basic skills (24 states), general knowledge (17 states), professional knowledge (20 states), and/or knowledge of the teacher's speciality area (21 states) (Anrig, Baratz-Snowden, & Goertz, 1987). Concomitantly, public policymakers have come to view the standardized test as the most widely accepted indicator of teaching effectiveness (Kubizyn & Borich, 1987). This phenomenon has occurred in spite of the absence of evidence that test scores predict successful teaching (Evertson, Hawley, & Zlotnik, 1984; Goodison, 1985).

Given, then, that teacher competency testing is a reality of professional life for educators at this time, considerable attention should be directed toward the selection of test content and the creation of a test paradigm more responsive to defining a sophisticated entry teacher profile. At the most fundamental level, teachers must know what to teach and how to teach it. Beyond such fundamental competence, teachers must know when to teach what and demonstrate a context-aware repertoire of teaching behaviors.

There is considerable consensus that a substantial knowledge base undergirds all teaching (Darlington-Hammond & Wise, 1983; Gideonse, 1989; Orlich, 1989; Saphier, 1985; Wittrock, 1987). Thus, a well-designed teacher competency test should address the knowledge base and incorporate what is known about effective teaching practices. In general, content should address such components as those identified by Shulman (1987): content knowledge; general pedagogical knowledge;

curriculum knowledge; pedagogical content knowledge; knowledge of learners and their characteristics; and knowledge of educational ends, purposes and values, and their philosophical and historical grounds (p. 3).

How this knowledge is assessed is of extreme importance. As the act of teaching itself is contextual, subject to a vast number of variables, so too is the testing experience. Paper and pencil tests alone are anemic representations of an individual's classroom performance. Such tests fail to capture the candidate's skill in fostering student learning. As presently constructed, there is considerable evidence that teacher competency testing may be biased against certain groups, particularly those greatly underrepresented as teachers in America's schools (Anrig, Baratz-Snowden, & Goertz, 1987).

Recently several states, notably Oregon, adopted policies for teacher licensure that incorporate evidence of pupil learning into licensure decisions (Schlack & Myton, 1988). In such cases, three major concerns are addressed: weak correlations between teacher behavior and pupil learning, significant variations in instructional assignments, and inadequate definitions of teacher competence.

The Educational Testing Service, publisher of the National Teacher Examinations, the National Professional Practices Board funded by the Carnegie Foundation, and Dr. Lee Shulman of Stanford University are working on models that will combine a number of approaches to teacher competency testing. The first two agencies seem committed to using paper and pencil tests in combination with problem solving opportunities via videotape analysis, some degree of observation of candidates, and

the analysis of a professional portfolio. Shulman's approach is attempting to create a new paradigm richly based on a massive research foundation.

It is appropriate at this point to mention a caveat related to the issue of who should design and identify appropriate content and format for teacher competency tests. To create a rich testing program with predictive validity requires the input of a variety of entities a priori rather than ex post facto. Schools and colleges of education, public school educators, and concerned citizens who are knowledgeable about pedagogy should have input into test design. Communication and collaboration among these groups should be fostered at both the grassroots and the highest levels. The data collected should then be both research-based and field-pertinent.

What then should be tested in entry-level teaching professionals? Certainly, testing should address candidates' awareness of and ability to apply the knowledge bases of teaching, content or discipline-based knowledge and ability to differentiate it according to students' needs and competences, interpersonal and oral and written communication skills, ability to reframe a problem by identifying new elements in it, and demonstration of the ability to plan and implement instruction so that students demonstrate measurable learning. A multiple assessment approach would create an elegant test combining written, interactive, and demonstration components. The written section should be criterion rather than norm-referenced. With such an approach, those concerned with the entry of qualified professionals to teach the nation's youth might then possess a competency testing model with greater validity than those currently in use.

Regardless of the type of competency examination utilized, paper and pencil or performance assessment, sound practice calls for validation studies to evaluate the correspondence between a given examination's actual characteristics and the stated purpose of the testing program. Using scores from competency examinations for certification decisions without conducting validity studies would constitute an improper use of the tests, according to the Equal Employment Opportunity Commission (1978), and would likely be challenged in court (Cross, 1985).

Validation Strategies

Teacher competency testing programs are usually designed to meet one or more of the following goals: (1) to assure that certified teachers have the knowledge and skills necessary to teach, (2) to assure that certified teachers act in a professionally acceptable manner, or (3) to assure that certified teachers positively and effectively influence students' development (Shulman, 1989). By law (e.g. United States v. South Carolina, 1979; Uniform Guidelines on Employee Selection Procedures, 1978; Standards for Educational and Psychological Testing, 1985), in-depth validation studies must be conducted to evaluate whether a test can provide a reasonable basis for making inferences concerning the goals. Ideally, these validation studies should involve gathering evidence that the test content is appropriate and gathering evidence that the test scores are related to desired outcomes.

This section discusses the methodological issues in the development and validation of teacher competency examinations. The procedures most often used in validation studies are described, some of the decisions

made in the process are identified, and issues stemming from these decisions are discussed.

Content Validity

According to the Standards for Educational and Psychological Testing, content validity evidence demonstrates the degree to which the sample of items or tasks on a test are representative of some defined domain of interest (American Psychological Association et al., 1985). However, as Yalow and Popham (1983) have noted, the "domain of interest" is not always apparent. Not surprisingly, most validation studies of teacher competency tests have provided diverse types of evidence to document the content validity of these tests.

Early validation studies provided evidence of the extent to which examinees would have an opportunity to learn the knowledge or skill tested by each item (Macmillan, 1982) and the extent to which the test specifications corresponded to the curricular coverage of teacher education programs (Radbil, 1983). Later studies have also provided evidence of the extent to which the content measured by each item is judged relevant to the domain of knowledge a teacher should possess (Masonis, 1987). In all cases, the evidence presented is in the form of professional judgments rendered by teacher educators, school teachers, or both. The above three types of content validity evidence may be referred to as (1) instructional validity, (2) program relevance validity, and (3) job relevance validity.

Instructional Validity. This validation strategy typically involves having faculty members from teacher education institutions judge whether or not a large percentage (e.g. 90%) of the students graduating from their

teacher education program would have an opportunity to learn the answer to each test item. Yalow and Popham (1983) have argued that evidence regarding examinee "opportunity to learn" may be important, but ought not be construed as evidence of content validity.

A more practical concern regarding instructional validity evidence is the question of whether any faculty member could be sufficiently well informed about the curriculum and what is actually taught in the teacher education program to know whether 90 percent of the students would have the opportunity to learn given knowledge. Consider, for example, what basis a professor of secondary social studies would have for knowing whether 90% of all education majors would have had an opportunity to learn the answers to questions found on the National Teacher Examinations Professional Knowledge Test involving topics as diverse as essentialist philosophy, motivation theory, Piaget's stages of cognitive development, or classroom assessments.

Program Relevance Validity. In this validation strategy, panels of teacher educators are provided test specifications and are asked to judge the degree of similarity between the topical emphasis in the test and the curriculum. As with instructional validity judgments, it would seem that faculty members would need to have far greater knowledge about the curriculum than most are likely to have. Even if panelists are provided advance copies of the test specifications and were to review the course syllabi for a given curriculum, it is doubtful whether judgments about the relative emphasis given each topic could be accurately made. The topic may not be explicitly referenced in any syllabus or the topic may appear in multiple syllabi with unknown emphasis.

Job Relevance Validity. In this validation strategy, panels of teacher educators and teachers are asked to judge the extent to which the content measured by each test item is relevant to the knowledge an entry-level teacher should possess in the domain measured by a particular test. A major limitation of this approach is that if there are knowledge areas critical to the job of teaching that are not measured by the test they will go undetected.

Another concern regarding job relevance validity is that, because job relatedness studies rely on expert judgment, results can differ depending on the sample surveyed. For example, in studying job relevence of a professional knowledge test, respondents with expertise in elementary school, middle school, high school, and special education have different perspectives which can be expected to affect their judgments. Differences between perspectives should be analyzed carefully, and validity studies should be planned with these perspectives in mind.

Teacher competency testing programs designed to assure that certified teachers engage in professionally acceptable practice are based on the inference that examinees with higher test scores are more likely to engage in professionally acceptable practice. It is debatable whether content validity evidence is a sufficient basis to justify the validity of the inference (Merwin, 1981). The argument in favor of additional evidence is that the inference is predictive: test scores should predict teacher performance in the classroom. Consequently, criterion-related validity evidence is required.

Criterion-Related Validity

Criterion-related validity is gathered to evaluate whether test scores are related to desired outcomes. There are two basic methods for collecting criterion-related validity evidence: concurrent and predictive. Several validation studies (Cross, 1985; Lovelace & Martin, 1984; Poggio et al., 1986) have focused on establishing concurrent validity by investigating the relationships between performance on teacher competency tests and success as an undergraduate student and performance on measures of academic achievement or scholastic aptitude.

No published studies, however, exist on the ability of scores from teacher competency examinations to predict actual classroom performance of either preservice or inservice teachers. State mandates calling for testing to certify "persons expected to be effective teachers" create the need for studies attempting to determine if teacher competency tests do indeed differentiate those judged as effective from those seen as ineffective. Though it seems logical that the greater the teacher's knowledge in the teaching field, the greater the chance of that person being an effective teacher, little research has been attempted which directly relates pedagogical and subject matter knowledge as measured by teacher competency tests and teacher classroom performance.

Although criterion-related validity evidence is clearly important for justifying some of the inferences made by using teacher competency tests, some argue that practical constraints make collecting the necessary data difficult. For example, it may not be possible to obtain criterion data for failing applicants barred from teaching positions. Also, not all

passing applicants take teaching positions; problems can arise if certificate holders who obtain teaching jobs are a nonrepresentative sample of those who pass the test. One method for confronting these problems is to collect criterion data during the last phase of the student teaching component of the teacher education program immediately prior to application for certification.

Another concern regarding criterion-related validity is the difficulty of obtaining valid criterion measures because of the wide variety of teaching contexts. Consequently, it is extremely difficult to determine criteria that apply to all examinees. There is nothing in the logic of criterion-related validity evidence, however, that requires a single study of all preservice teachers seeking certification. Moreover, it may be possible to separate the teaching contexts into subsets (i.e., early childhood, elementary, middle school, high school, special education) that are sufficiently homogeneous for common criteria to apply and to collect predictive evidence in several of these subsets.

Recommendations Regarding Validity Evidence

Given the central role played by validation in all testing activities, validation is an especially important issue for teacher competency tests because of profound changes in the world of teaching. These changes in how teachers teach and how their profession will develop influence the nature of the inferences that we draw from teacher competency test scores. Largely as a result of social forces that extend well beyond the world of education, the emphasis of teacher assessment is on the licensing function served by such assessments.

Particularly in this complex environment, it is critical to consider what are appropriate validation strategies. To consider this topic in the correct context, we must move away from a point of view that conceptualizes different "types" of validity, all of which are separate entities having their own discrete methods and purposes, toward a view that validity is a summary judgment made on the basis of numerous interrelated pieces of evidence. First, given the current purpose of teacher competency examinations to demonstrate knowledge and skills necessary for beginning teachers, content validity is an absolute necessity. It must be established that the test items are relevant to the work of teaching and that the test content adequately and accurately represents the pedagogical and subject matter domains it purports to sample.

A second interrelated piece of evidence critical to the consideration of appropriate validation strategies addresses the predictive inferences made from teacher competency examinations. Given state mandates calling for testing to certify "persons expected to be effective teachers," predictive validity evidence is an absolute necessity. It must be established that teacher competency tests do indeed differentiate those judged as effective from those seen as ineffective.

Finally, appropriate validation strategies should be concerned in the future with the heretofore unconsidered issue of the construct validity of teacher competency tests. Inferences about test scores should be clearly related to psychological and educational theory about what constitutes effective teachers and teaching. Conducting construct validation studies, however, exposes judgments made on the basis of

theoretical and philosophical values and, as such, thrusts the question of validity into legal and social arenas.

It is the view of Samuel Messick (1987) that the consequential basis of validity which implies a concern for fairness is an essential element of the validation of any test. This is a critical concern for competency tests because no matter what form of assessment we use for teachers, issues of adverse impact and intrinsic fairness of the tests will have social and legal significance. Standardization and the accompanying legal and ethical considerations, therefore, become issues in the development and use of teacher competency examinations.

Standardization

All examinees seeking certification in the same jurisdiction should face comparable testing conditions. The threat of denial of certification makes it imperative to control as many relevant variables as possible to avoid advantaging some examinees and penalizing others. In an effort to control the variables necessary to assure fairness and consistency, most teacher certification agencies have turned to standardized written and/or performance examinations.

This section discusses legal and equity issues that confront teacher testing activities. A brief overview of litigation involving claims against teacher testing programs is provided, impacts of teacher testing on minority groups are described, and philosophical and practical considerations emerging from equity issues are addressed.

Legal Considerations

The wave of teacher certification testing programs has been accompanied by legal challenges to certification and employment decisions predicated on

the results of this testing. Standardized tests are frequently accused of being discriminatory because a disproportionate percentage of minority examinees perform poorly. Therefore, the use of these tests as a prerequisite to public employment has been challenged on the basis of the equal protection clause of the Fourteenth Amendment that prohibits states from denying citizens equal protection.

The Supreme Court has rejected such equal protection claims, however, if it can be shown that the use of the test is rationally related to a legitimate employment objective and not accompanied by a discriminatory motive. In the leading case, National Education Association v. South Carolina (1977), the Supreme Court affirmed a lower court's conclusion that South Carolina's use of the National Teacher Examinations for teacher certification purposes satisfied Fourteenth Amendment guarantees since the test had a rational relationship to the legitimate purpose of improving the effectiveness of the state's teaching force and was not administered with discriminatory intent. The trial court was satisfied that the test was valid in that it measured knowledge of the course content in teacher preparation programs and was justified by the legitimate employment objective of assuring that teachers possessed basic knowledge.

In 1986, the Fifth Circuit Court of Appeals also overturned a preliminary injunction against using a basic skills test as a prerequisite to enrollment in teacher education programs in Texas (United States v. Lulac, 1986). The court reasoned that a state is not obligated to educate or certify teachers who cannot pass a valid test of basic skills necessary for professional training.

Title VII of the Civil Rights Act of 1964 has also served as a legal rationale for challenges to certification based on standardized tests. Title VII prohibits employment discrimination on the basis of race, color, religion, sex, or national origin. The law applies to public or private employers with 15 or more employees, employment agencies, and labor organizations. Where facially neutral requirements, such as a test, have a disparate impact on a group protected by Title VII, the employer must establish a business necessity for the challenged practice.

Through several decisions, the Supreme Court elaborated on the business necessity standard under Title VII. In Griggs v. Duke Power Company (1971), the Court concluded that the employer must substantiate that tests used as a condition of employment are related to job performance in order to satisfy Title VII. In a subsequent decision (Albemarle Paper Company v. Moody, 1975), the Court recognized that Title VII requires an employment test to be validated for the specific jobs for which it is used.

In 1978, the Equal Employment Opportunity Commission (EEOC) issued Uniform Guidelines on Employee Selection Procedures designed to assure that employment practices with an adverse impact on a group protected by Title VII are justified by a business necessity. Under the guidelines, employment tests are expected to be (1) reliable in that the measurement instrument is accurate and provides dependable data, and (2) valid in that the instrument actually measures what it purports to measure. The EEOC prefers test validation by correlating job performance with test scores but will allow other validation methods that satisfy professionally recognized standards.

Significant precedents, therefore, have been established in the area of employment testing in general and teacher testing in particular. Despite these legal efforts to insure fairness in testing practices and the use of test results, however, there remains considerable evidence that current teacher competency tests are not equally valid and fair for all examinees.

Equity Issues

The use of standardized tests as competency indices for teacher certification is challenging an important societal issue: equity. With the implementation of teacher testing programs, there has been a drop in the supply of talented, well-educated minority teachers. As a result, a host of philosophical and practical issues has emerged which must be addressed every time teacher competency testing is discussed.

According to the Commission on Teacher Credentialing (1984), of 6,644 minority candidates in California who took the California Basic Educational Skills Test (CBEST) in 1983, 3,854, or 58 percent, failed. The highest failure rate was among Blacks. Of the 2,040 Blacks who took the test, only 530, 26 percent, were able to proceed with their plans to become teachers. Furthermore, only 39 percent (834) of 2,133 Hispanics passed and only 50 percent (637) of 1,259 Asians passed the CBEST. In comparison, the pass rate for Whites was 76 percent, with 18,856 of the 24,540 Whites passing.

The pass rates elsewhere are no more encouraging. On Florida's Teacher Competency Examination, while 90 percent of the White candidates passed the examination, only 35 percent of Black candidates, 51 percent

of Hispanic candidates, and 63 percent of Asian candidates passed (Smith, 1984). Passing rates for the Georgia Certification Tests were comparable: 89 percent of White first-time test takers passed the test while only 34 percent of Blacks and 48 percent of Hispanics did (Goertz, Ekstrom, & Coley, 1984). The first administration of the Texas testing program for prospective teachers eliminated 84 percent of the Black candidates and 65 percent of the Hispanic candidates on the basis of mathematics and reading examinations (Smith, 1984).

The use of standardized tests as competency indices for teacher certification also impacts equity with regard to candidates with handicapping conditions. While advance testing arrangements can be made with most test publishers and/or testing sites to accommodate the special needs of individuals with specific physical disabilities such as visual or hearing impairments, a dyslectic faces unique problems as a result of potentially inflexible testing situations. While an argument could be made that such an individual lacks essential skills to present effective instruction, the reality is that many individuals entering the field of special education do so to have a positive impact within a system that may have been less than compassionate to them.

The problems associated with high failure rates among minority groups are made all the more serious by the need for qualified minority and special education teachers. Of the nation's 35 largest city school districts, the majority now have overwhelmingly minority-student enrollments (Gallegos, 1988). Furthermore, elementary and secondary schools across the nation have experienced a 34 percent increase in the

enrollment of students with handicapping conditions as defined by the provisions of Public Law 94-142 (Williams, 1987). When teacher competency testing is considered in the context of these trends, it appears, as Airasian (1987) argues, that "the crucial issues of testing are not technical....(but) social, economic, and value-laden" (p. 411).

Conclusions and Recommendations

Teacher competency testing is not a fleeting phenomenon but rather a response to the burgeoning reform movement of the 1980's in American education. Despite the substantial evidence that standardized tests are not infallible indicators of teaching ability, more than 40 states presently rely on them for initial teacher licensure.

Clearly, there must be vigilance in insuring the proper use of standardized tests in assessing the competency of potential teachers. Test developers must be persuaded to revise and design tests that are both unbiased and valid. Where necessary, legal action should be considered as an avenue for test reform and policy change. There must also be full understanding of what tests can and cannot do. To quote William Raspberry (1983), tests cannot

> ...measure patience, love of children and learning, and a hundred other things that make up teacher competency. But the tests can measure whether a teacher has learned the basics of pedagogic techniques (which we consider important, else why would we mandate education courses for teachers?) and whether a teacher has solid grasp of the material to be taught... (p. 37)

While tests cannot be relied upon to identify who has the personal warmth, drive, and dedication required of a good teacher, they can provide some reliable information about basic pedagogical and subject matter knowledge. Further, more than measuring personal aptitude, good competency tests indicate what potential teachers have learned; ideally, they also show how potential teachers are able to apply their learning to what the tests ask of them.

The future generation of competency examinations seem to presage some combination of pencil-and-paper testing of one or more knowledge bases, problem solving through simulation or interactive video, and a professional dossier or portfolio. As such, a notable difference in the new multi-method form of assessment from the former exclusively pencil-and-paper examinations is that the testing process will be labor intensive and costly. Who will do the actual administration and scoring of certification tests and who will bear the costs of testing are yet to be definitively determined.

Unquestionably, the content of teacher competency tests, regardless of format, will have a powerful impact on the nature of teacher education curriculum nationwide. The phenomenon has already been demonstrated in Texas and other states which issue study guides for certification tests. The time seems ripe for teacher education institutions to assume a leadership role in defining the nature, content, and purpose of teacher competency testing <u>a priori</u> rather than after the fact. With this approach, teacher education curriculum and teacher competency testing will enjoy a symbiotic relationship. Colleges of teacher education will become involved not when the testing is done but at the time the

groundwork for the testing process is laid. Thus, the teacher education profession could insure that what is taught is appropriate to the needs of an entry level professional and that what is tested is actually what is taught in teacher preparation programs.

Further, the 1980's have heralded a substantial decentralization movement in organizations from the federal government through major private and public corporations. Given the American tradition of local autonomy over education, it would seem appropriate that teacher competency testing, rather than relying on a single nationwide model, would evolve into vehicles for testing sound generally accepted pedagogical principles without neglecting locally designated curricular emphases.

Certainly, there are challenges to be met with regard to teacher competency testing. Crucial decisions in these grave professional matters must be made judiciously in a variety of areas but there should be no fear of seeking alternate methods of qualifying those who will teach the youth of this nation. Perhaps teacher competency testing has been accepted too soon and too unequivocably as a panacea without fully recognizing the problems associated with teacher testing programs and without adequately identifying other modes of assessing novice professional competence.

References

Airasian, P. (1987). State-mandated testing and educational reform: Context and consequences. American Journal of Education, 95, 393-412.

Albemarle Power Company v. Moody. (1975). 422 U.S. 401, 432

American Psychological Association, American Educational Research Association, & Council on Measurement in Education. (1985). Standards for educational and psychological tests. Washington, DC: Author.

Anrig, G., Baratz-Snowden, J., & Goertz, M. (1987). Testing policies and minority participation in teaching: A look at the research. Paper presented to the annual meeting of the American Educational Research Association, Washington, DC.

California Commission on Teacher Credentialing. (1984). CBEST performance in relation to personal background factors. Sacramento: Author.

Cross, L. (1985). Validation of the NTE tests for certification decisions. Educational Measurement: Issues and Practices, 27, 7-10.

Darlington-Hammond, L., Wise, A., & Pease, S. (1983). Teacher evaluation in the organizational context: A review of the literature. Review of Educational Research, 53, 285-328.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978, August). Uniform guidelines on employee selection procedures. Federal Register, 43, 38290-38315.

Evertson, C., Hawley, W., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. Journal of Teacher Education, 36, 2-12.

Gallegos, A. (1988). The negative consequences of teacher competency testing. Phi Delta Kappan, 69, 631.

Gideonse, H. (1989). Relating knowledge to teacher education: Responding to NCATE's knowledge base and related standards. Washington, DC: American Association of Colleges of Teacher Education.

Goertz, M., Ekstrom, R., & Coley, R. (1984). The impact of state policy on entrance into the teaching profession. Final report to the National Institute of Education, NIE-G-83-0073. Princeton, NJ: Educational Testing Service.

Griggs v. Duke Power Company. (1971). 401 U.S. 424

Kubizyn, T., & Borich, G. (1987). Educational testing and measurement. Glenview, IL: Scott Foresman.

Lovelace, T., & Martin, C. (1984). The revised NTE as a predictor of teachers' performance in public school classrooms. Lafayette, LA: University of Southeastern Louisiana. (ERIC Document Reproduction Service No. ED 351 416)

Macmillan, J. (1982). Content validation study of the NTE by 5-12 teachers and instructional supervisors. Richmond, VA: Virginia Commonwealth University and Educational Testing Service.

Masonis, E. (1987). Report on a study of selected NTE and GRE tests by the state of New Jersey. Princeton, NJ: Educational Testing Service.

Merwin, J. (1981). Review of the NTE. In A. Buros (Ed.), The eighth mental measurements yearbook (Vol. I). Highland Park, NJ: Gyphon.

Messick, S. (1987). Validity. Princeton, NJ: Educational Testing Service.

Murray, S. (1986). Synthesizing teacher testing policy options. Portland, OR: Northwest Regional Educational Laboratory.

National Education Association v. South Carolina. (1977). 445 F. Supp. 1094 (DSC), aff'd, 434 U.S. 1026 (1978)

Orlich, D. (1989). Education reforms: Mistakes, misconceptions, miscues. Phi Delta Kappan, 70, 512-517.

Poggio, J., Glasnapp, D., Miller, N., Toliefson, N., & Berry, J. (1986). Strategies for validating teacher certification tests. Educational Measurement: Issues and Practices, 28, 18-26.

Radbil, A. (1983). Report on a study of the NTE Core Battery by the state of New Mexico. Princeton, NJ: Educational Testing Service.

Raspberry, W. (1983). From analysis to policy. In R. Bratt, Testing teachers. Berkeley: University of California Policy Analysis for California Education.

Rudner, L. (Ed.). (1987). What's happening in teacher testing: An analysis of state teacher testing practices. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Sandefur, J. (1984, August). Standards for admissions to teacher education. Issue paper prepared for the Minnesota Higher Education Coordinating Board.

Schalock, H., & Myton, D. (1988). A new paradigm for teacher licensure: Oregon's demand for evidence of success in fostering learning. Journal of Teacher Education, 39, 8-16.

Shulman, L. (1987). Assessment for teaching: An initiative for the profession. Phi Delta Kappan, 68, 38-44.

Shulman, L. (1989). The paradox of teacher assessment. In New directions for teacher assessment. Princeton, NJ: Educational Testing Service.

Shulman, L., & Sykes, G. (1986). A national board of teaching? In search of a bold standard. Washington, DC: Carnegie Forum on Education and the Economy.

Smith, G. (1984). The critical issue of excellence and equity in competency testing. Journal of Teacher Education, 35, 7.

Sweeney, J., & Maratt, R. (1986). Teacher evaluation. In R. Berk (Ed.), Performance assessment: Methods and applications. Englewood Cliffs, NJ: Prentice-Hall.

United States v. Lulac. (1986). 628 F. Supp. 304 (E.D. Texas), rev'd, 793 F. 2d. 636 (5th Cir.)

United States vs. South Carolina. (1977). 445 F. Supp. 1094

Williams, A. (1987). The impact of PL 94-142 on school enrollment patterns. Journal of Special Education, 38, 162-170.

Yalow, E., & Popham, W. (1983). Content validity at the crossroads. Educational Researcher, 12, 10-14, 21.